Pisco Med Publishing

Original Research Article

An Improved HS Type Spectral Conjugate Gradient Method

Jing Li, Shujie Jing

School of Mathematics and Information Science, Henan Polytechnic University, Henan, China

Abstract: The spectral conjugate gradient (SCG) method was one of the most commonly used method to solve nonlinear unconstrained optimization problems. In this paper, an improved HS spectral conjugate gradient method is proposed. Under common assumptions, the descent and global convergence of this method are proved.

Keywords: Unconstrained optimization; Global convergence; Line search; Sufficient descend conditions

1. Introduction

We consider the unconstrained optimization problem: min $f(x)$,where $f:R^n \rightarrow R$ is continuously differentiable and bounded, $X_0 \in R^n$ is an initial solution, give the iterative formula of spectral conjugate gradient method: $X_{k+1}=x_k+\alpha_k d_k$ $k\geq 1$, the step factor β_k is a positive parameter, the search direction d_k is defined by $d_k = \begin{cases} -g_k, & k=1, \\ -\theta_k g_k + \beta_k d_{k-1}, & k>1. \end{cases}$

where g_k is the gradient of f at a point x_k ,the scalar β_k is the conjugate direction control parameter. Different selection of β_k corresponds to different the conjugate gradient(CG) methods. Many scholars have studied the convergence of the CG method under different line search. Some authors use accurate line search to calculate the step size, and some authors use imprecise line search to calculate the step size, such as strong Wolfe line search, which is defined as follows.

$$d_k = \begin{cases} -g_k, & k=1, \\ -\theta_k g_k + \beta_k d_{k-1}, & k>1. \end{cases} \quad (1) \qquad u_k - \big\lfloor -\theta_k g_k + \beta_k d_{k-1}, k>1. \quad (2) \qquad d_k = \begin{cases} -g_k, & k=1, \\ -\theta_k g_k + \beta_k d_{k-1}, & k>1. \end{cases} \quad (3)$$

Aiguo et al[1] proposed an improved FR type spectral conjugate gradient method ; Chen et al[2] proposed a new PRP type spectral conjugate gradient method ; Neculai [3] proposed a three-term conjugate gradient algorithm ; Wang et al [4] proposed two new HS type spectral conjugate gradient methods; Fang et al[5] proposed a modified DY spectrum conjugate gradient method ; Jing et al[6] constructed a new spectral conjugate gradient method by using Armijo line search; Li et al [7] proposed an improved spectral conjugate gradient method, the search direction generated by iteration is fully reduced and globally convergent.Inspired by the literature [8, 9], a new conjugate parameters and spectral parameters are proposed, in which the search direction is usually as follows

Multiply the left and right ends of equation (2) by y_{k-1} to o btain $d_k^T y_{k-1}=-\theta_k g_k^T y_{k-1}+\beta_k^{LF} d^T_{k-1} y_{k-1}$(4)

We can get from[10] $d_k=-B_k g_k$, $d^T_k y_{k-1}=-(B_k g_k)^T y_{k-1}=-g^T_k S_{k-1}$,. The above relationship means that if the line search is accurate, the above formula holds. At this time, $g_k^T S_{k-1}=0$, the actual numerical algorithm usually adopts imprecise line search instead of accurate line search. Therefore, use condition $d_k^T y_{k-1}=-t g_k^T S_{k-1}$. It seems more reasonable to replace the conjugate condition, substitute into formula (4) $d_k^T y_{k-1} = -\theta_k g_k^T y_{k-1} + \beta_k^{LF} d_{k-1}^T y_{k-1}$, Divide by $g_k^T y_{k-1}$,

$$\theta_k = \frac{t g_k^T s_{k-1}}{g_k^T y_{k-1}} + \frac{d_{k-1}^T y_{k-1}}{\left|g_k^T d_{k-1}\right| - \mu d_{k-1}^T g_{k-1}} \text{ Because of } d_k^T y_{k-1} = -\theta_k g_k^T y_{k-1} + \beta_k^{LF} d_{k-1}^T y_{k-1} \text{, so } \theta_k = \frac{t g_k^T \alpha_k d_{k-1}}{g_k^T y_{k-1}} + \frac{d_{k-1}^T y_{k-1}}{\left|g_k^T d_{k-1}\right| - \mu d_{k-1}^T g_{k-1}} \quad (5)$$

2. New Algorithm and its descent

Algorithm 2.1

Step1. Given an initial value $x_0 \in R^n$, parameter ε>0, 0≤μ≤1, t>0. Step2.Set k=0, $d_0=-g_0$. If $\|g_0\| < \varepsilon$, stop

Step3.Determine the step length α_k satisfying (1) Step4. Calculate x_{k+1}, if $\|g_0\| < \varepsilon$, stop.

Step5. Use (3), (2), and (5) to calculate , and respectively Step6. $k=k+1$ Turn Step3.

Lemma 2.1 Assuming that the search direction of the SCG method is (2), β_k^{LF} is defined by (3), and step α_k is obtained by (1), the

doi: 10.18282/l-e.v10i3.2422

following sufficient descent conditions are satisfied $g_k^T d_k = -\eta \|g_k\|^2$ (6)

Proof: Using mathematical induction . If k=0, then the search direction d_k is given by d_k=-g_k ,so there is d_0=-g_0, $g_k^T d_k = g_0^T d_0 = -\|g_0\|^2$,Therefore, condition (6) is true, assuming that $k \geq 0$ is true, condition (6) is also true. By multiplying g_k^T both sides of equation (2), we can get the following equation

$$g_k^T d_k = -\left(\frac{t g_k{}^T \alpha_k d_{k-1}}{g_k{}^T y_{k-1}} + \frac{d_{k-1}^T y_{k-1}}{|g_k^T d_{k-1}| - \mu d_{k-1}^T g_{k-1}} \right) \|g_k\|^2 + \frac{d_{k-1}^T y_{k-1}}{|g_k^T d_{k-1}| - \mu d_{k-1}^T g_{k-1}} \|g_k^T\|^2 = -\frac{t g_k{}^T \alpha_k d_{k-1}}{g_k{}^T y_{k-1}} \|g_k\|^2$$

Let $\eta = \frac{t g_k{}^T \alpha_k d_{k-1}}{g_k{}^T y_{k-1}}$, we can get $g_k^T d_k = -\eta \|g_k\|^2$ .This method satisfies sufficient descent condition for any k .

Lemma 2.2 Let the sequences $\{d_k\}$ be generated by algorithm 2.1,then for any $k \geq 0$, the relation $\beta_k^{LF} \leq \frac{\|g_k\|^2}{|g_k^T d_{k-1}|}$ holds.

Proof : Because of $y_{k-1} = g_k - g_{k-1}$, we get the $g_k^T y_{k-1} = g_k^T (g_k - g_{k-1}) = \|g_k\|^2 - g_k^T g_{k-1}$, so we can get the formula

$$\beta_k^{LF} = \frac{g_k^T y_{k-1}}{|g_k^T d_{k-1}| - \mu d_{k-1}^T g_{k-1}} \leq \frac{g_k^T y_{k-1}}{|g_k^T d_{k-1}|} = \frac{g_k^T (g_k - g_{k-1})}{|g_k^T d_{k-1}|} \leq \frac{\|g_k\|^2}{|g_k^T d_{k-1}|}.$$

3. Global Convergence

Assumption (A)

1. f(x)has a lower bound on the level set $S = \{x \in R^n \,|\, f(x) \leq f(x_0)\}$, where x_0 is the initial iteration point.

2. f(x)is continuously differentiable in a certain neighborhood N of the level set S , and the gradient function g(x) satisfies Lipschitz continuity, there is a constant $L > 0$ which makes $\|g(x) - g(y)\| \leq L\|x - y\|, \forall x, y \in N$ hold.

Lemma 3.1 If Assumption A holds, the search direction d_k satisfies $g_k d_k$, then sequences $\{g_k\}$ and $\{d_k\}$ generated by algorithm 2.1 satisfy Zoutendijk condition, then $\sum_{k \geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty$ holds.

Theorem 3.1 If Assumption A holds, the sequence $\{g_k\}$ generated by algorithm 2.1.Then $\liminf_{k \to \infty} \|g_k\| = 0$.(7)

Proof: Assume that (7) does not hold.There must be a constant $\gamma > 0$, make that $\|g_k\| \geq \gamma, k \geq 0$. Derived from formula (2),

$d_k + \theta_k g_k = \beta_k^{LF} d_{k-1}$, Square both sides at the same time, $(d_k + \theta_k g_k)^T (d_k + \theta_k g_k) = (\beta_k^{LF})^2 \|d_{k-1}\|^2$, Expanded

$\|d_k\|^2 + 2\theta_k d_k^T g_k + \theta_k^2 \|g_k\|^2 = (\beta_k^{LF})^2 \|d_{k-1}\|^2$,Transposition of terms $\|d_k\|^2 = (\beta_k^{LF})^2 \|d_{k-1}\|^2 - 2\theta_k d_k^T g_k - \theta_k^2 \|g_k\|^2$,divide both sides by $(g_k^T d_k)^2$ at the same time. Because of $\beta_k^{LF} \leq \frac{\|g_k\|^2}{|g_k^T d_{k-1}|}$, $g_k d_{k-1} = y_{k-1}^T d_{k-1} + d_{k-1}^T g_{k-1}$ So

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{\|g_k\|^4}{(g_k^T d_{k-1})^2} \cdot \frac{\|d_{k-1}\|^2}{\eta^2 \|g_k\|^4} - \frac{2\theta_k}{\eta \|g_k\|^2} - \frac{\theta_k^2}{\eta^2 \|g_k\|^2} \leq \frac{\|d_{k-1}\|^2}{\eta^2 (g_{k-1}^T d_{k-1})^2} + \frac{1}{\|g_k\|^2} = \frac{\|d_{k-1}\|^2}{\eta^4 \|g_{k-1}\|^4} + \frac{1}{\|g_k\|^2}$$

, When $k = 0$, $\|d_0\|^2 / (g_0^T d_0)^2 = 1 / \|g_0\|^2$ .It can be obtained the above inequality $\|d_k\|^2 / (g_k^T d_k)^2 \leq \sum_{i=1}^{k} 1 / \eta^4 \|g_i\|^2 \leq k / \eta^4 \gamma^2$

Therefore $(g_k^T d_k)^2 / \|d_k\|^2 \geq \eta^4 \gamma^2 / k = +\infty$ .Contradictions with lemma 3.1 above $\liminf_{k \to \infty} \|g_k\| = 0$.

References:

[1] Aiguo Lu et al. A variant spectral-type FR conjugate gradient method and its global convergence[J]. Applied Mathematics and Computation, 2011, 217(12) : 5547-5552.

[2] Chen LW, Xia F Q, Jia Z Y. A new PRP type spectral conjugate gradient method and its global convergence [J].Journal of He- nan Normal University (Natural Science),2013,41(03): 5-9+14.

[3] Neculai Andrei. A new three-term conjugate gradient algorithm for unconstrained optimization[J].Numerical Algorithms,2015, 68(2):305-321.

[4] Wang G F, Shan R, Huang W, et al. Two newspectralconjugategradient algorithms based on Hestenes - Stiefel.J. Algorithms Comput. Technol. 2017,11(4):345–352.[5]Fang M L, Sun M, A modified DY spectral conjugate gradient method based on Wolfe line searching, Journal of Yangtze University(Nat Scl Edit),15(13)(2018) 7-10.